多参数PET/CT影像组学联合ResNet-50深度学习预测非小细胞肺癌病理亚型

何健平[1],郭 鹏[2],夏露花[2],李俊鑫[1],雷楷笛[2],王新华[2]
[1]新疆医科大学第三临床医学院,新疆 乌鲁木齐 830011;[2]新疆医科大学附属肿瘤医院核医学科,新疆 乌鲁木齐 830011

摘要:目的 探索基于深度学习的多参数PET/CT影像组学在术前无创预测非小细胞肺癌(NSCLC)病理亚型中的临床价值。**方法** 回顾性收集2018年7月~2024年12月于新疆医科大学附属肿瘤医院接受^{18}F-FDG PET/CT检查,且经病理证实为NSCLC的患者279例(腺癌187例、鳞癌92例),按7:3随机分为训练集(*n*=195)与验证集(*n*=84)。使用3DSlicer完成肿瘤分割,组内相关系数评估勾画一致性;PyRadiomics提取PET及CT手工影像组学特征,ResNet-50提取深度特征;分别以逻辑回归、支持向量机、K近邻和ExtraTrees构建临床模型、影像组学模型和深度学习模型,选取最优模型结合临床危险因素构建联合预测模型。采用受试者工作特征曲线、校准曲线及决策曲线评估模型的预测效能和临床获益。**结果** 多因素logistic回归分析显示,年龄、吸烟史、病灶位置、SUV_{max}及CYFRA21-1是区分肺腺癌与鳞癌的独立预测因子。临床模型、CT影像组学模型、PET影像组学模型、CT深度学习模型、PET深度学习模型及综合模型的训练集AUC分别为0.941、0.894、0.865、0.882、0.879、0.958;验证集AUC分别为0.922、0.792、0.868、0.848、0.873、0.948。决策曲线结果显示,综合模型的临床净获益更优。**结论** 融合临床因素的^{18}F-FDG PET/CT深度学习影像组学模型在术前无创地预测NSCLC病理亚型方面有很好的应用价值。
关键词:深度学习;影像组学;非小细胞肺癌;^{18}F-FDG PET/CT

Multivariate PET/CT imageomics combined with ResNet-50 deep learning for predicting pathological subtypes of non-small cell lung cancer

HE Jianping[1], GUO Peng[2], XIA Luhua[2], LI Junxin[1], LEI Kaidi[2], WANG Xinhua[2]
[1]Third Clinical Medical College of Xinjiang Medical University, Urumqi 830011, China; [2]Department of Nuclear Medicine, Affiliated Tumor Hospital of Xinjiang Medical University, Urumqi 830011, China

Abstract: Objective To investigate the clinical utility of deep learning-based multiparametric ^{18}F-FDG PET/CT radiomics for preoperative non-invasive prediction of histological subtypes in non-small cell lung cancer (NSCLC). **Methods** This retrospective study included 279 patients with pathologically confirmed NSCLC (187 adenocarcinomas and 92 squamous cell carcinomas) who underwent ^{18}F-FDG PET/CT at the Affiliated Cancer Hospital of Xinjiang Medical University from July 2018 to December 2024. The cohort was randomly assigned to training (*n*=195) and validation (*n*=84) sets at a 7:3 ratio. Tumor segmentation was performed using 3DSlicer, and segmentation consistency was evaluated using the intraclass correlation coefficient. Handcrafted radiomics features were extracted from PET and CT images using PyRadiomics, whereas deep learning features were extracted using a ResNet-50 architecture. Clinical, radiomics, and deep learning models were constructed using logistic regression, support vector machine, K-nearest neighbors, and ExtraTrees classifiers. The optimal model was subsequently integrated with clinical risk factors to develop a combined prediction model. Discrimination, calibration, and clinical utility were assessed using ROC curves, calibration curves, and decision curve analysis. **Results** Multivariate logistic regression analysis revealed that age, smoking history, lesion location, maximum standardized uptake value, and CYFRA21-1 were independent predictors for distinguishing lung adenocarcinoma from squamous cell carcinoma. In the training cohort, the clinical model, CT radiomics model, PET radiomics model, CT deep learning model, PET deep learning model, and combined model achieved areas under the curve (AUCs) of 0.941, 0.894, 0.865, 0.882, 0.879, and 0.958, respectively; the corresponding AUCs in the validation cohort were 0.922, 0.792, 0.868, 0.848, 0.873, and 0.948. Decision curve analysis indicated that the combined model provided a superior clinical net benefit. **Conclusion** Integrating clinical factors with ^{18}F-FDG PET/CT deep learning radiomics enables accurate preoperative non-invasive prediction of NSCLC histological subtypes and holds considerable promise for clinical application.
Keywords: deep learning; radiomics; non-small-cell lung cancer; ^{18}F-FDG PET/CT

肺癌是全球发病率与死亡率最高的恶性肿瘤,2022年全球肺癌死亡病例约达187万例[1]。非小细胞癌(NSCLC)约占肺癌总病例的85%[2],其主要的2个病理类型为腺癌和鳞癌,腺癌好发于女性,多无吸烟史,病灶较小,早期症状隐匿;鳞癌则以男性吸烟者为主,病灶较大,症状明显。两者在分子特征、药物反应及预后方面差异显著[3]。因此,明确NSCLC不同亚型对制定个体化治疗方案具有重要指导意义[4]。

目前临床主要依赖侵入性穿刺活检明确病理类型,但小标本难以反映肿瘤整体异质性,且存在出血、气胸等并发症,直接影响诊疗方案的制定及预后评估的准确

收稿日期:2026-01-08
基金项目:国家高性能医疗器械创新中心开放基金项目(NMED2024CX-03-007)
作者简介:何健平,在读硕士研究生,E-mail: hejianping1998@163.com
通信作者:王新华,主任医师,E-mail: xj1w@sina.com

性[5]。^{18}F-FDG PET/CT作为无创检查技术,可同步提供全身多器官分子功能与解剖信息,已成为肺癌分期、疗效评估和复发监测的重要工具[6]。近年来,影像组学技术通过高通量提取医学影像定量特征[7],为NSCLC病理亚型鉴别提供了新途径。既往研究表明,基于CT或PET的影像组学模型在区分腺癌与鳞癌方面展现出良好效能:有学者基于CT影像组学构建的模型验证集AUC达0.893[8];有研究联合临床特征与PET影像组学,最优模型AUC为0.859[9];也有研究融合临床指标与PET/CT影像组学特征,联合模型验证集AUC提升至0.901[10]。然而,上述多数研究仅单独使用CT或PET影像组学,未能充分利用双模态互补优势;传统影像组学依赖人工预设特征,对图像深层信息的挖掘能力有限;现有联合模型性能仍有提升空间,且缺乏对深度学习特征的系统性探索。

深度学习技术通过多层卷积神经网络自动学习图像层次化特征,可捕捉人眼难以识别的细微纹理差异[11]。有研究比较了多种传统机器学习模型与VGG16深度学习网络,结果显示VGG16模型AUC达0.903,显著优于传统方法[12]。有研究进一步证实,融合CT影像组学与深度学习特征可将分类AUC提升至0.944[13]。然而,目前深度学习在NSCLC病理亚型预测中的应用仍处于探索阶段,现有研究多基于单模态CT图像,PET/CT双模态深度特征提取及与影像组学的融合策略尚不充分;如何有效整合临床危险因素、影像组学特征与深度学习特征构建多参数综合模型,以实现预测效能的最大化,仍需进一步验证。

基于此,本研究拟构建融合临床特征、PET/CT手工影像组学特征及ResNet-50深度学习特征的多参数机器学习模型,构建"临床-影像组学-深度学习"三维度的综合预测体系,旨在系统比较传统影像组学与深度学习的效能差异,探索PET/CT双模态特征融合策略,并验证多参数联合模型的增益效应,为术前无创精准预测NSCLC病理亚型提供新的技术路径。

1 资料与方法

1.1 研究对象

本研究回顾性选取2018年7月~2024年12月于新疆医科大学附属肿瘤医院核医学科接受^{18}F-FDG PET/CT检查且经病理确诊为NSCLC的患者。本研究经我院伦理委员会批准(审批号:伦科批字[2024]第271号)。所有患者免除知情同意。纳入标准:年龄≥18岁;术后或穿刺活检病理证实的腺癌和鳞癌;^{18}F-FDG PET/CT检查均于术前或穿刺前1月内完成;SUV_{max}>2.50,病灶直径>1.00 cm。排除标准:有抗肿瘤治疗史;临床病理数据记录不完整;图像质量不佳,难以准确勾画病灶VOI;既往有恶性肿瘤病史。最终纳入279例患者,其中腺癌187例,鳞癌92例,收集患者的临床信息和^{18}F-FDG PET/CT影像学资料。本研究采用分层随机抽样方法,以病理类型为分层依据,按照7:3比例将279例患者划分为训练集与验证集,其中训练集195例(包括腺癌129例,鳞癌66例),验证集84例(包括腺癌58例,鳞癌26例)。

1.2 检查仪器与图像分析

所有患者均采用Philips Ingenuity TF PET/CT扫描仪检查。^{18}F-FDG显像剂,核素由日本住友HM-10HC回旋加速器生产,标记合成借助北京派特化学合成系统完成,放射化学纯度检测结果均>95%。检查前嘱患者空腹6 h以上,确保空腹血糖值<11.0 mmol/L;以3.70~5.55 MBq/kg的剂量静脉注射显像剂,待患者静卧45~60 min达到显像剂摄取平衡后,取仰卧位进行PET/CT采集。CT扫描参数:管电压120 kV,管电流300 mA,螺距0.705,矩阵512×512;PET三维采集,1 min/床位,矩阵144×144。常规扫描后加1.0 mm薄层肺CT扫描,扫描时保持屏气。图像经迭代算法重建,获取全身多平面断层图像,传输至Medex工作站后处理。PET/CT图像由1位初级职称和1位高级职称核医学医师双盲分析,诊断分歧时通过讨论达成共识。

1.3 肿瘤分割与预处理

通过MedEX工作站提取所有患者的5 mm层厚PET与1 mm层厚胸部CT的DICOM格式图像,经批量格式转换获得NIfTI-1文件;继而将其载入开源医学影像分析平台3DSlicer5.8.1实施肺肿瘤病灶标注。标注过程采用盲法设计,低年资医师分别在PET图像(阈值40%)与CT图像(窗宽1200 HU、窗位-600 HU)上,逐层手动勾勒肿瘤的感兴趣区(ROI);高年资医师再对勾画结果进行复核校正,最终达成标注结果的一致性要求,ROI勾画示例(图1)。鉴于图像异质性可能对后续影像组学分析造成干扰,本研究在特征提取环节前,对原始PET与CT图像实施标准化预处理流程。先使用灰度离散化将原始灰度值重映射至统一区间,实现跨图像的对比度与亮度校准;随后采用线性插值将图像重采样为1 mm×1 mm×1 mm的各向同性体素,消除像素尺寸差异,确保所有体素对应一致的物理空间尺度,最大限度减少个体差异带来的偏倚。

1.4 特征提取与筛选

使用PyRadiomics3.0.1软件提取影像组学特征。提取的特征体系涵盖3大核心类别,即形状特征、一阶统计特征及纹理特征;其中纹理特征具体包含灰度依赖矩阵(GLDM)、灰度共生矩阵(GLCM)、灰度区域大小矩阵(GLSZM)、灰度游程长度矩阵(GLRLM)以及邻域灰度差值矩阵(NGTDM)5大亚类[14]。特征预处理与筛

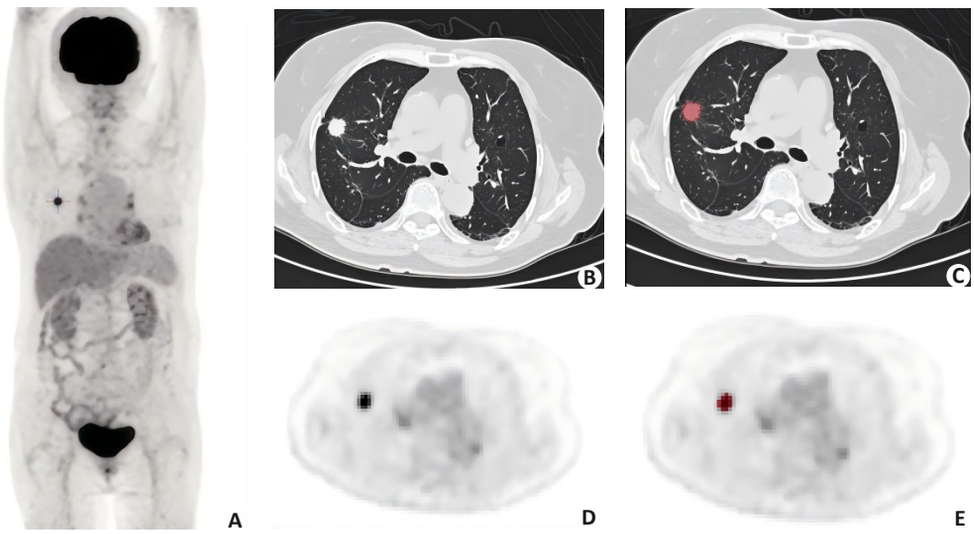

图1 感兴趣区勾画过程

Fig. 1 Region of interest delineation process. **A:** Whole-body PET image. **B, D:** CT and PET tomographic images; **C, E;** Images after delineation.

选流程包括：首先对各影像组学特征进行*Z*-score标准化处理，以此消除：不同特征间数值范围差异所造成的干扰，再筛选组内相关系数（ICC）>0.75的影像组学特征，采用独立样本*t*检验或Mann-Whitney *U*检验验证筛选后特征在两种组织学亚型间的差异。通过Pearson相关系数识别并剔除高度相关特征（舍弃平均绝对相关系数>0.9的特征），减少特征冗余与噪声，避免模型过拟合。最后采用10折交叉验证的最小绝对收缩和选择算子（LASSO）算法筛选出非零系数的特征变量。

提取深度学习特征采用在ImageNet数据集上预训练的ResNet-50卷积神经网络作为基准模型，通过迁移学习优化模型适配性，基于交叉熵损失函数完成端到端训练，实现输入至输出的全流程自主优化[15]。本研究依据影像组学勾画的ROI定位病灶，提取包含肿瘤最大横截面的单一层级作为网络输入，以保留肿瘤形态与代谢异质性信息；所有入选切片经重采样统一为224×224分辨率，本策略以关键层面表征替代完整三维上下文建模，未引入相邻层面的空间关联。模型训练采用批量大小为32的小批量梯度下降，选用Adam优化器，初始学习率0.001，配合余弦退火调度器历经100个epoch逐步衰减，以促进稳定收敛并防范过拟合。数据增强环节实施随机水平翻转、±15°平面旋转、0.9~1.1倍随机缩放及弹性形变等在线扩增，提升模型对体位差异及扫描变异的鲁棒性。数据集按患者级别实施分层随机划分，以7:3比例分配至训练集与内部验证集，确保同一受试者的全部切片仅出现于单一子集，杜绝数据泄露。

1.5 模型构建及效能评价

基于筛选所得的临床特征、影像组学特征和深度学习特征，通过逻辑回归（LR）、支持向量机（SVM）、K近邻（KNN）和极端随机树（ExtraTrees）4种机器学习算法构建临床模型、CT影像组学模型、PET影像组学模型、CT深度学习模型、PET深度学习模型，最终采用决策层融合策略[16]，整合深度学习模型、影像组学模型与临床模型，构建多参数综合模型，实现对患者组织学亚型的全面预测。模型诊断效能的量化评估指标包括：ROC曲线下面积（AUC）及其95% *CI*、准确度、敏感度、特异度、阳性预测值（PPV）、阴性预测值（NPV）及F1分数。同时通过校准检验评估模型拟合优度；采用决策曲线分析（DCA）评估该预测模型在临床场景中的净获益，验证其实际应用价值。

1.6 统计学分析

采用Python3.7.12进行统计学分析。符合正态分布的计量资料以均数±标准差表示，组间比较根据方差齐性选择Student *t*检验或Mann-Whitney *U*检验；计数资料以*n*（%）表示，组间差异比较采用卡方检验或Fisher精确概率法。通过ROC曲线及AUC值评估模型预测性能。以*P*<0.05为差异有统计学意义。

2 结果

2.1 临床模型构建

2.1.1 资料特征分析 在训练集与验证集中，年龄、性别、吸烟史、肿瘤位置、CYFRA21-1、SCCA、SUV_{max}及TLG在腺癌与鳞癌中的差异有统计学意义（*P*<0.05，表1）。

2.1.2 临床独立预测因素分析 单因素和多因素logistic回归分析结果显示，年龄、吸烟史、肿瘤位置、SUV_{max}和CYFRA21-1是区分ADC和鳞癌的独立危险因素（表2）。基于上述特征构建临床预测模型后，训练集与验证集中，LR模型的AUC为0.875、0.937，SVM模型的AUC为0.792、0.885，KNN模型的AUC为0.899、0.814，ExtraTrees模型的AUC为0.941、0.922。ExtraTrees模型的预测效能最优。

表1 训练集和验证集NSCLC患者的临床资料比较

Tab.1 Comparison of clinical data of NSCLC patients in the training set and the validation set

Clinical factors	Training set (*n*=195)		*P*	Validation set (*n*=84)		*P*
	ADC	SCC		ADC	SCC	
Age (year, *Mean*±*SD*)	61.69±11.05	67.14±9.73	<0.001	62.90±8.53	71.35±10.19	<0.001
Gender [*n*(%)]			<0.001			<0.001
Male	64(49.61)	59(89.39)		24(41.38)	24(92.31)	
Female	65(50.39)	7(10.61)		34(58.62)	2(7.69)	
Smoking [*n*(%)]			<0.001			<0.001
No	83(64.34)	15(22.73)		42(72.41)	5(19.23)	
Yes	46(35.66)	51(77.27)		16(27.59)	21(80.77)	
Tumor location [*n*(%)]			<0.001			0.006
Peripheral	107(82.95)	32(48.48)		49(84.48)	14(53.85)	
Central	22(17.05)	34(51.52)		9(15.52)	12(46.15)	
Spiculation [*n*(%)]			<0.001			0.465
No	28(21.71)	34(51.52)		14(24.14)	9(34.62)	
Yes	101(78.29)	32(48.48)		44(75.86)	17(65.38)	
Lobulation [*n*(%)]			0.405			0.569
No	14(10.85)	4(6.06)		6(10.34)	1(3.85)	
Yes	115(89.15)	62(93.94)		52(89.66)	25(96.15)	
Pleural retraction [*n*(%)]			0.117			0.444
No	9(6.98)	10(15.15)		6(10.34)	5(19.23)	
Yes	120(93.02)	56(84.85)		52(89.66)	21(80.77)	
CYFRA21-1 (ng/mL, *Mean*±*SD*)	86.83±58.87	139.13±57.94	<0.001	78.31±52.92	154.38±44.88	<0.001
CEA (μg/L, *Mean*±*SD*)	123.95±73.67	128.43±66.60	0.679	122.52±78.12	119.08±72.01	0.849
CA125 (U/mL, *Mean*±*SD*)	106.41±67.54	100.43±57.29	0.539	93.10±64.64	95.73±60.46	0.861
SCCA (ng/mL, *Mean*±*SD*)	1.58±5.02	4.59±5.80	<0.001	0.88±0.51	3.42±4.77	<0.001
ProGRP (pg/mL, *Mean*±*SD*)	38.31±16.69	38.43±17.34	0.965	37.50±11.63	112.15±351.82	0.108
NSE (ng/mL, *Mean*±*SD*)	15.26±8.17	16.18±6.88	0.434	15.08±7.03	18.02±13.83	0.199
SUV_{mean} (*Mean*±*SD*)	4.03±1.78	6.24±2.23	<0.001	4.44±4.98	6.17±1.62	0.088
SUV_{max} (*Mean*±*SD*)	10.79±6.28	17.15±6.64	<0.001	10.04±5.70	17.37±4.75	<0.001
TLG (*Mean*±*SD*)	112.01±148.05	645.96±909.07	<0.001	108.23±146.87	369.59±395.68	<0.001

CYFRA21-1: Cytokeratin fragment antigen 21-1; CEA: Carcinoembryonic antigen; CA125: Cancer antigen 125; SCCA: Squamous cell carcinoma antigen; ProGRP: Pro-gastrin-releasing peptide; NSE: Neuron-specific enolase; SUV_{max}: Maximum standardized uptake value; SUV_{mean}: Standardized uptake value mean; TLG: Total lesion glycolysis; ADC: Adenocarcinoma; SCC: Squamous cell carcinoma.

表2 单因素及多因素分析logistic回归分析显示可用于鉴别腺癌与鳞癌的临床特征

Tab.2 Univariate and multivariate logistic regression analyses revealed the clinical characteristics that could be used to differentiate ADC from SCC

Clinical factors	Univariate analysis		Multivariate analysis	
	OR (95% *CI*)	*P*	OR (95% *CI*)	*P*
Age	1.012(1.008,1.017)	<0.001	1.006(1.002,1.010)	0.006
Smoking	1.491(1.370,1.623)	<0.001	1.209(1.088,1.343)	0.003
Tumor location	1.447(1.313,1.595)	<0.001	1.188(1.084,1.302)	0.002
SUV_{max}	1.032(1.025,1.038)	<0.001	1.010(1.002,1.017)	0.044
CYFRA21-1	1.003(1.003,1.004)	<0.001	1.001(1.001,1.002)	0.001

2.2 影像组学模型构建

经特征筛选后，获得16项最优CT影像组学特征，包括4项一阶特征和12项纹理特征，纹理特征中包含2项GLCM、7项GLSZM、1项GLRLM和2项NGTDM；15项最优PET影像组学特征，包含2项形状特征、3项一阶特征和10项纹理特征，纹理特征中包含2项GLCM、5项GLSZM、1项GLRLM和2项GLDM，分别构建CT与PET影像组学模型。CT影像组学模型在训练集与验证集的AUC值为LR模型：0.894、0.792；SVM模型：0.889、0.758；KNN模型：0.871、0.777；ExtraTrees模型：0.898、0.785；综合来看，LR模型的预测效能最优。PET影像组学模型训练集与验证集中AUC为LR模型：0.894、0.811；SVM模型：0.874、0.826；KNN模型：0.865、0.868；ExtraTrees模型：0.867、0.841；其中，KNN模型的预测效能最优。

2.3 深度学习模型构建

深度学习模型训练完成后，从ResNet-50倒数第2层（全连接层前的特征层）提取高维图像特征，采用主成分分析将特征维度降至64维，降低特征冗余的同时提升计算效率，将降维后的图像特征与影像组学特征进行多模态融合，最终筛选得到18项最优CT深度学习特征，包括4项一阶特征、12项纹理特征、2项深度学习特征；纹理特征中包含2项GLCM、3项GLRLM、6项GLSZM、1项GLDM。17项最优PET深度学习特征，包括1项形状特征、2项一阶特征、11项纹理特征、3项深度学习特征，纹理特征中包含1项GLCM、1项GLRLM、5项GLSZM、4项GLDM。分别构建CT深度学习模型与PET深度学习模型。CT深度学习模型训练集与验证集中AUC为LR模型：0.894、0.790；SVM模型：0.853、0.834；KNN模型：0.890、0.783；ExtraTrees模型：0.882、0.848；其中，ExtraTrees模型的预测效能最优。PET深度学习模型训练集与验证集中AUC为LR模型：0.885、0.858；SVM模型：0.893、0.851；KNN模型：0.879、0.873；ExtraTrees模型：0.889、0.855。综合来看，KNN模型的预测效能最优。

2.4 综合模型构建及各模型效能评估

2.4.1 综合模型构建 将临床模型、影像组学模型与深度学习模型进行整合，采用逻辑回归算法构建多模态联合预测模型，联合模型列线图（图2）。CYFRA21-1为鳞癌特异性标志物，评分跨度最大、权重最高，其水平上升显著提升鳞癌判定概率。年龄评分随龄递增，符合鳞癌老年高发特征。SUV_{max}反向计分，代谢活性越高分值越大，结果显示鳞癌葡萄糖摄取强于腺癌。病灶部位及吸烟史为二分类变量，中央型病变和吸烟阳性均获最高评分，结果显示鳞癌中央气道起源及烟草暴露相关。影像组学模型输出为0~1连续型变量，系LASSO筛选特征综合输出，线性刻度与风险正相关。CT深度特征概率于0.1~1.0区间密集分布，残差网络高维特征在中等风险区极敏感；PET深度模型概率呈0~1线性均匀分布，代谢组特征呈稳定线性贡献。

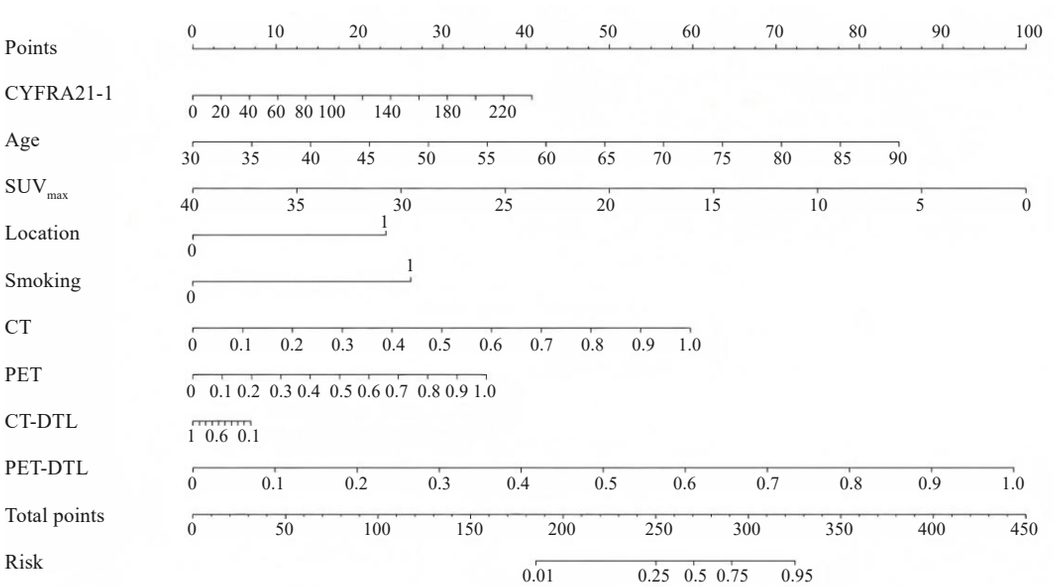

图2 联合模型鉴别腺癌与鳞癌的列线图

Fig.2 Nomogram of the combined model for differentiating ADC from SCC.

综合模型在训练集与验证集中均展现出优异的鉴别效能（表3）。在训练集中，AUC为0.958（95% *CI*=0.930~0.987），准确度和敏感度分别为0.933和0.833。在验证集中，该模型的AUC为0.948（95% *CI*=0.903~0.994），准确度为0.905，敏感度为0.885（图3），均优于单一类型模型，具有较好的预测性能。

表3 6种模型在训练集和验证集上的效能比较

Tab.3 Comparison of the performance of 6 models on the training set and the validation set

Dataset	Model	AUC (95% *CI*)	Accuracy	Sensitivity	Specificity	PPV	NPV	F1
Training set	CT	0.894(0.850-0.938)	0.826	0.773	0.853	0.729	0.880	0.750
	PET	0.865(0.817-0.914)	0.779	0.894	0.721	0.621	0.930	0.733
	CT-DL	0.882(0.834 -0.931)	0.800	0.773	0.814	0.680	0.875	0.723
	PET-DL	0.879(0.831-0.927)	0.779	0.833	0.752	0.632	0.898	0.719
	Clinical	0.941(0.909-0.973)	0.892	0.803	0.938	0.869	0.903	0.835
	Combined	0.958(0.930-0.987)	0.933	0.833	0.984	0.965	0.920	0.894
Validation set	CT	0.792(0.696-0.889)	0.679	0.962	0.561	0.490	0.970	0.649
	PET	0.868(0.793 -0.944)	0.798	0.731	0.828	0.655	0.873	0.691
	CT-DL	0.848(0.764 -0.931)	0.786	0.885	0.741	0.605	0.935	0.719
	PET-DL	0.873(0.799-0.947)	0.786	0.808	0.789	0.618	0.900	0.700
	Clinical	0.922(0.864-0.980)	0.857	0.885	0.845	0.719	0.942	0.793
	Combined	0.948(0.903 -0.994)	0.905	0.885	0.914	0.821	0.946	0.852

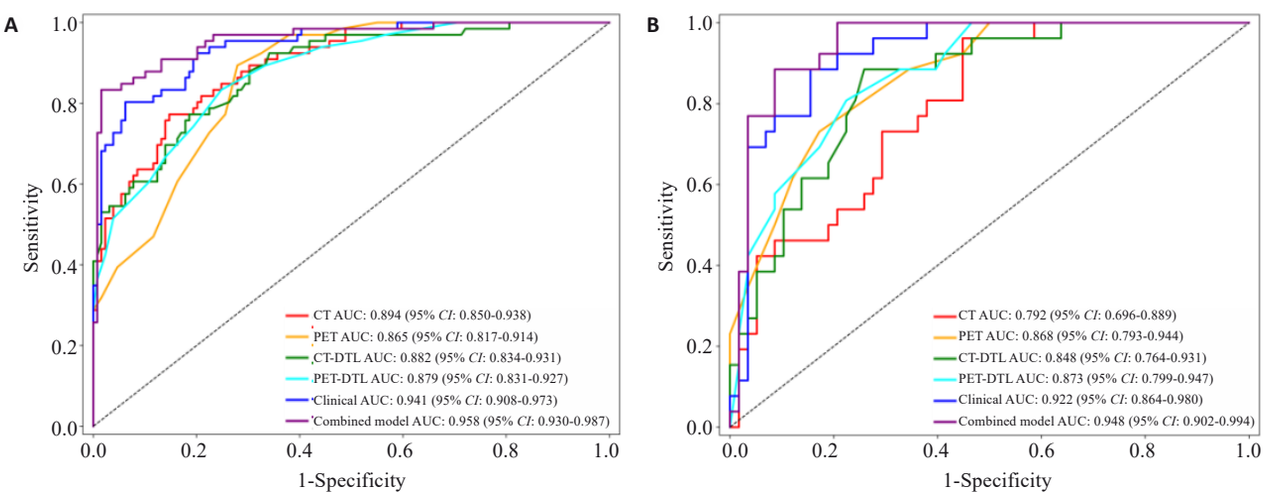

图3 临床、影像组学、深度学习模及联合模型的ROC曲线对比

Fig. 3 Comparison of ROC curves of clinical, radiomics, deep learning models and combined models. **A:** Training set; **B:** Validation set.

2.4.2 模型决策曲线及校准曲线分析 决策曲线结果显示，联合模型在0.1~0.9的阈值概率区间内均保持较高净获益，优于各单模态模型，具有良好的临床应用价值（图4）。校准曲线显示，联合模型预测概率与实际观测概率的拟合线紧密贴近理想对角线，模型具有良好的校准度和可靠性（图5）。

3 讨论

肺癌仍是全球首要癌症负担，其中NSCLC占主导地位。准确识别NSCLC病理亚型是实施个体化治疗、提升生存的关键。本研究以LR、SVM、KNN、ExtraTrees四类算法分别建立临床、影像组学及深度学习模型，均取得不错的判别能力；而融合临床因素的联合模型在训练与验证集的AUC分别达0.958、0.948，性能表现更佳。

本研究发现年龄、吸烟史、肿瘤位置、SUV_{max}及CYFRA21-1指标对区分腺癌和鳞癌具有统计学意义。本研究纳入患者多为老年人，鳞癌患者年龄大于腺癌患者，这与既往研究[17]报道一致，可能与老年患者烟龄长、免疫功能低下有关[18]。吸烟患者以鳞癌为主要类型，与既往研究[19]相符。肿瘤位置方面，腺癌更常见于外周部位，而鳞癌多位于中央部位，这与既往研究[20]一致。血清CYFRA21-1水平在鳞癌中显著高于腺癌[21]。SUV_{max}作为临床最常用评估病灶代谢活性的半定量指标，可反映出病灶组织最活跃部位的葡萄糖代谢水平[22]，本研究结果显示鳞癌患者中SUV_{max}值明显高于腺癌患者，可能与鳞癌中葡萄糖转运蛋白表达水平更高有关[23]。而本研究中传统CT形态学特征，如分叶征、毛刺征及胸膜牵拉，在两亚型间差异无统计学意义，主观判定偏差可

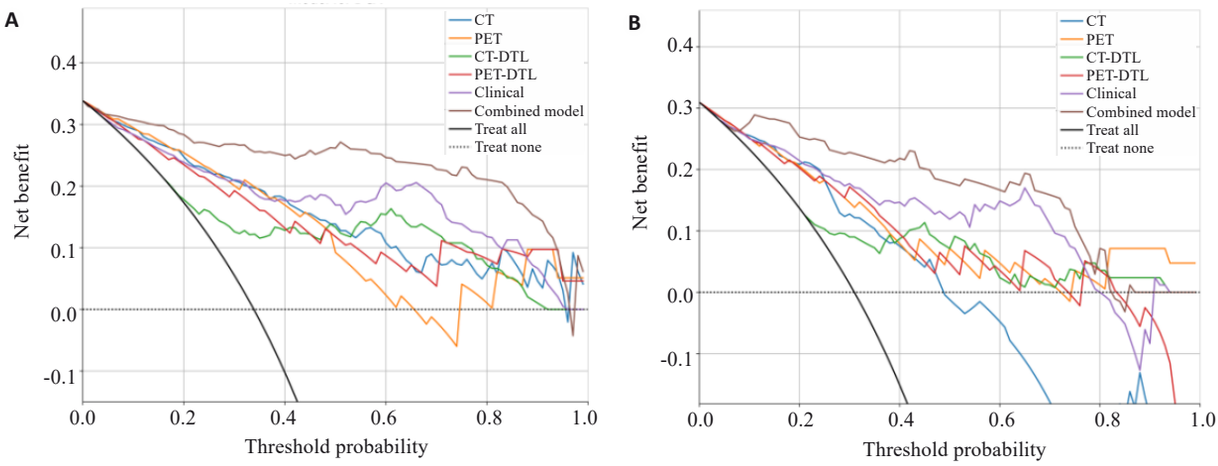

图4 临床、影像组学、深度学习及联合模型的决策曲线对比

Fig.4 Comparison of decision curves of clinical, radiomics, deep learning and combined models. **A:** Training set; **B:** Validation set.

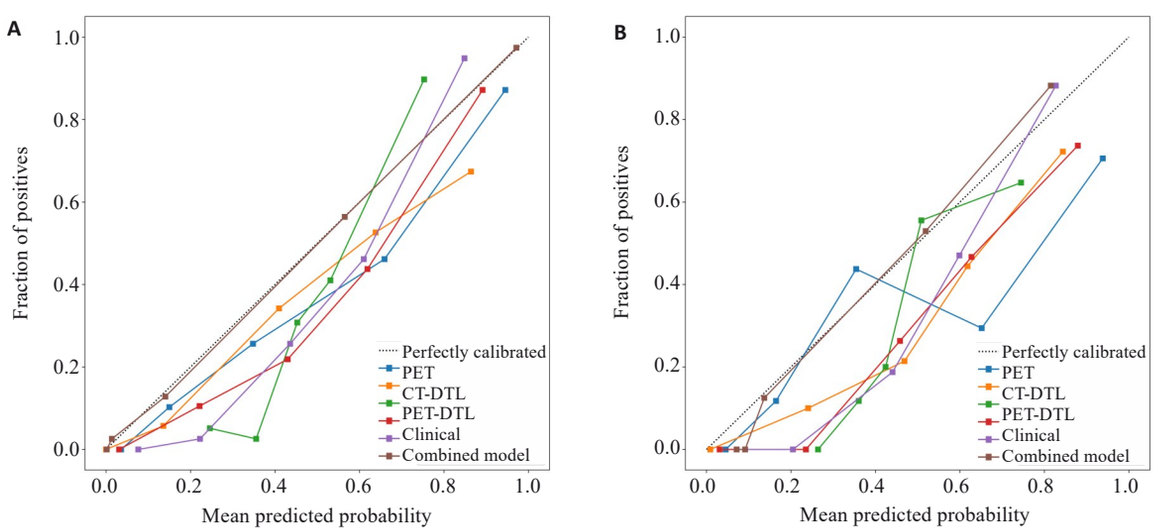

图5 临床、影像组学、深度学习及联合模型的校准曲线对比

Fig.5 Comparison of calibration curves for clinical, radiomics, deep learning, and combined models. **A:** Training set; **B:** Validation set.

能是判别效能受限的主因[24]。

影像组学通过深入分析影像图像，提取具有临床价值的参数，量化图像异质性以间接反映肿瘤细胞异质性[7]。在NSCLC组织学分类领域，多项研究已证明其潜力和有效性。有研究收集了129例腺癌和鳞癌患者CT图像，通过LASSO逻辑回归从129例患者中筛选出5项关键特征（含一阶特征、GLCM及GLRLM纹理特征），模型训练集AUC达0.905，验证集AUC为0.893[8]。有团队基于CT图像构建的4分类（腺癌、鳞癌、大细胞癌及未另行指定）模型在独立验证集实现86%准确率，同时证实一阶特征、GLCM、GLSZM 3种特征更有利于NSCLC亚型的分类[25]。本研究中的CT影像组学模型专注灰度纹理特征，与上述研究结果相同。有研究分析81例NSCLC患者PET/CT图像，提取了8项PET和8项CT一阶特征，发现鳞癌体积、摄取及异质性均高于腺癌[26]。本研究PET影像组学模型除了一阶特征，还系统提取了形状特征和多类纹理特征，更加全面地量化肿瘤的表型特征。

传统机器学习算法包括LR、SVM、KNN、Extra-Trees、随机森林等。有研究提取了396例NSCLC患者（210例腺癌、186例鳞癌）的4项临床特征和40项PET组学特征，在随机森林、神经网络、朴素贝叶斯、LR、SVM 5种机器学习算法中，LR模型性能最佳，AUC达0.859[9]。本研究中LR算法构建的模型性能表现均比较优秀，对处理多参数数据展现出较高的稳健输出与精准预测能力。

深度学习通过端到端框架将原始数据映射至高阶表征，全程无需人工设计特征，显著提升了模型对异质性图像的解析精度与效率[27]。有研究利用867例ADC和552例鳞癌患者的PET/CT图像，系统比较了10种特

筛选方法、10种经典机器学习模型及VGG16深度网络。结果显示，线性判别分析（LDA）结合SVM获得最佳传统性能，验证集AUC达0.863，准确率约0.79；而VGG16进一步将AUC提升至0.903，准确率增至0.841，优于所有传统模型[12]。本研究结果与该研究相符，深度学习模型在本研究中同样展现出优于或相当于传统影像组学模型的性能。另有研究指出，联合深度特征与经典影像组学可再度提升模型区分度。有学者基于CT影像组学和深度学习区分337例NSCLC患者的病理亚型，应用7种滤波预处理并设计3D CNN、3D ResNet-18/101三支深度学习网络，最终融合模型AUC达0.89[28]。上述证据提示，混合特征策略优于单一范式，本研究亦遵循该思路。

本研究除了对CT图像和PET图像的影像组学特征进行分析，还结合了临床特征。有研究将临床指标与^{18}F-FDG PET/CT影像组学特征融合构建区别腺癌与鳞癌的二分类机器学习模型，其联合模型在训练集AUC为0.932，验证集AUC为0.901，DeLong检验及决策曲线均显著优于单一子模型[10]。综合上述优势，本研究提出的融合模型在NSCLC病理亚型判别精度上优于既往文献报道。

本研究存在以下局限：尽管本研究构建的模型在训练集上展现出较优的性能，但在部分测试集上的表现有所回落，说明模型存在过拟合风险，扩大样本量有望缓解该问题；本研究为单中心回顾性分析，缺乏外部队列验证，限制了结果的可重复性与临床推广价值，未来需建立独立验证集，并纳入跨种族、跨地域的多中心数据，系统评估模型在不同人群及设备参数下的稳定性；数据集中ADC比例高于SCC，存在类别不平衡，同时仅纳入NSCLC两大常见亚型，未覆盖小细胞肺癌及其他罕见亚型，削弱了模型的全面性，后续将构建包含多亚型的大样本队列，探索更具普适性的多分类框架；本研究仅对比4种传统机器学习与1种深度学习算法，尽管已获理想性能，但算法选择与超参数优化仍有提升空间，随着人工智能快速迭代，可进一步引入Transformer、自监督学习、图神经网络等精细架构，并整合基因组学与临床病理多维特征，构建高表达能力的多模态融合模型，以提升预测精度与生物学可解释性。

综上所述，本研究构建的融合模型在同类研究中展现出优异的区分效能，可无创、精准地鉴别腺癌与鳞癌，为个体化治疗与预后评估提供可靠依据，并辅助临床决策优化。

参考文献：

[1] Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024[J]. CA A Cancer J Clin, 2024, 74(1): 12-49.

[2] Nicholson AG, Tsao MS, Beasley MB, et al. The 2021 WHO classification of lung tumors: impact of advances since 2015[J]. J Thorac Oncol, 2022, 17(3): 362-87.

[3] Ettinger DS, Wood DE, Aisner DL, et al. Non-small cell lung cancer, version 3.2022, NCCN clinical practice guidelines in oncology[J]. J Natl Compr Cancer Netw JNCCN, 2022, 20(5): 497-530.

[4] Imyanitov EN, Iyevleva AG, Levchenko EV. Molecular testing and targeted therapy for non-small cell lung cancer: Current status and perspectives[J]. Crit Rev Oncol, 2021, 157: 103194.

[5] Ren F, Fei Q, Qiu K, et al. Liquid biopsy techniques and lung cancer: diagnosis, monitoring and evaluation[J]. J Exp Clin Cancer Res, 2024, 43(1): 96.

[6] Sauter AW, Schwenzer N, Divine MR, et al. Image-derived biomarkers and multimodal imaging strategies for lung cancer management[J]. Eur J Nucl Med Mol Imaging, 2015, 42(4): 634-43.

[7] Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: Extracting more information from medical images using advanced feature analysis[J]. Eur J Cancer, 2012, 48(4): 441-6.

[8] Zhu XZ, Dong D, Chen ZD, et al. Radiomic signature as a diagnostic factor for histologic subtype classification of non-small cell lung cancer[J]. Eur Radiol, 2018, 28(7): 2772-8.

[9] Hyun SH, Ahn MS, Koh YW, et al. A machine-learning approach using PET-based radiomics to predict the histological subtypes of lung cancer[J]. Clin Nucl Med, 2019, 44(12): 956-60.

[10] Ren CY, Zhang JP, Qi M, et al. Machine learning based on clinico-biological features integrated 18F-FDG PET/CT radiomics for distinguishing squamous cell carcinoma from adenocarcinoma of lung[J]. Eur J Nucl Med Mol Imaging, 2021, 48(5): 1538-49.

[11] LeCun Y, Bengio Y, Hinton G. Deep learning[J]. Nature, 2015, 521 (7553): 436-44.

[12] Han Y, Ma Y, Wu ZY, et al. Histologic subtype classification of non-small cell lung cancer using PET/CT images[J]. Eur J Nucl Med Mol Imaging, 2021, 48(2): 350-60.

[13] Panchawagh S, Halder A, Haldule S, et al. Quantitative radiomic analysis of computed tomography scans using machine and deep learning techniques accurately predicts histological subtypes of non-small cell lung cancer: a retrospective analysis[J]. Eur J Surg Oncol, 2025, 51(10): 110376.

[14] Takehana K, Sakamoto R, Fujimoto K, et al. Peritumoral radiomics features on preoperative thin-slice CT images can predict the spread through air spaces of lung adenocarcinoma[J]. Sci Rep, 2022, 12: 10323.

[15] Vutukuri M, Habibullah PS. Automated classification of lung cancer subtypes cells using microscopic images and ensembled deep learning architectures[J]. Sci Rep, 2026, 16: 235.

[16] Chou YB, Hsu CH, Chen WS, et al. Deep learning and ensemble stacking technique for differentiating polypoidal choroidal vasculopathy from neovascular age-related macular degeneration [J]. Sci Rep, 2021, 11: 7130.

[17] Wang BY, Huang JY, Chen HC, et al. The comparison between adenocarcinoma and squamous cell carcinoma in lung cancer patients[J]. J Cancer Res Clin Oncol, 2020, 146(1): 43-52.

[18] Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36

cancers in 185 countries[J]. CA Cancer J Clin, 2021, 71(3): 209-49.

[19] 王成弟, 陈勃江, 宋璐佳, 等. 23228例肺癌患者临床流行病学及病理特征趋势分析[J]. 华西医学, 2020, 35(7): 813-20.

[20] Xie XQ, Li XL, Tang WJ, et al. Primary tumor location in lung cancer: the evaluation and administration[J]. Chin Med J, 2021, 135 (2): 127-36.

[21] 戴 倩, 王 梦, 黄 钢. 基于^{18}F-FDG PET/CT影像组学的非小细胞肺癌病理亚型分类[J]. 中国医学物理学杂志, 2023, 40(4): 416-22.

[22] Kocaman G, Ibrahımov F, Kahya Y, et al. SUV_{max} of the lesion should be considered in the treatment plan for stage I non-small cell lung cancer[J]. Ann Nucl Med, 2025, 39(8): 792-8.

[23] Schuurbiers OCJ, Meijer TWH, Kaanders JHAM, et al. Glucose metabolism in NSCLC is histology-specific and diverges the prognostic potential of 18FDG-PET for adenocarcinoma and squamous cell carcinoma[J]. J Thorac Oncol, 2014, 9(10): 1485-93.

[24] 胡峙珩, 赵艳萍, 巴 雅, 等. 基于^{18}F-FDG PET/CT影像组学特征和临床因素预测肺腺癌EGFR突变状态[J]. 中国医学影像学杂志, 2023, 31(8): 844-51.

[25] Liu J, Cui JJ, Liu F, et al. Multi-subtype classification model for non-small cell lung cancer based on radiomics: SLS model[J]. Med Phys, 2019, 46(7): 3091-100.

[26] Bianconi F, Palumbo I, Fravolini ML, et al. Texture analysis on [^{18}F] FDG PET/CT in non-small-cell lung cancer: correlations between PET features, CT features, and histological types[J]. Mol Imaging Biol, 2019, 21(6): 1200-9.

[27] Jiang YH, Yang M, Wang SH, et al. Emerging role of deep learning-based artificial intelligence in tumor pathology[J]. Cancer Commun, 2020, 40(4): 154-66.

[28] Liang BY, Tong C, Nong JY, et al. Histological subtype classification of non-small cell lung cancer with radiomics and 3D convolutional neural networks[J]. J Imaging Inform Med, 2024, 37 (6): 2895-909.

（编辑:郎 朗）